

Multiagent RL Optimal Execution

Andrea Macrì, Fabrizio Lillo
Scuola Normale Superiore

May 19, 2024

Abstract

Optimal Execution with two agents.

Contents

1	Introduction	1
2	Model	1
2.1	Market Model	1

1 Introduction

In this paper we will consider the problem of optimal execution in a simple market environment with *two agents*. The agents aim at liquidating an initial inventory q_0 of stocks S over N steps within a time window $[0, T]$. The problem, in its single agent formulation, has been at the center of a lively stream of research. In fact, starting from the seminal works of Almgren and Chriss [**almgren**] and Bertsimas and Lo [**bertsimas**], many other contributions have expanded on the problem and proposed different solutions, either by changing the market impact functional form or by considering other market features, notable contributions in this area are Guéant and Lehalle (2015); Gueant, Lehalle, and Fernandez-Tapia (2012); Cartea and Jaimungal (2016); Casgrain and Jaimungal (2019); Cartea, Jaimungal, and Penalva (2015); Bouchaud, Farmer, and Lillo (2009); Bouchaud et al. (2003); Gatheral, Schied, and Slynko (2012); Obizhaeva and Wang (2013).

One interesting approach is to use reinforcement learning (RL), where the agent learns the optimal strategy by interacting with the environment. For example, Ning, Lin, and Jaimungal (2021) used a Double Deep Q-network to find the best execution schedule in a market with temporary impact using real stock data. Jeong and Kim (2019) developed a trading strategy for various stock indexes using a Deep Q-network. Dab’erius, Granat, and Karlsson (2019) compared the performance of DDQN and Proximal Policy Optimization (PPO) algorithms on artificial data and found that both outperformed the TWAP strategy in suboptimal scenarios and performed similarly to a slice-and-dice strategy in optimal scenarios. Finally, Macrì and Lillo (2024) used DDQN to show how RL is empirically able to learn optimal trading policies when liquidity is time-varying.

2 Model

In this paper we will consider the problem of optimal execution in a simple market environment with *two agents*. The agents aim at liquidating an initial inventory q_0 of stocks S over N steps within a time window $[0, T]$. We divide the time window $[0, T]$ into N time-steps such that: $t = 1, \dots, N$ and $\Delta t = T/N$. The price of the stock at time t is denoted by P_t . The agents can only trade at the mid-price P_t and the market impact is modeled as a linear function of the trading rate. The market impact is given by:

2.1 Market Model